Solving Schrodinger equation with soft constrained monotonic neural network

Han-Lin Li[1], Xuan-liu[1], Kai-Fang Pu[1], Long-Gang Pang[2]

(1. College of Science, Wuhan University of Science and Technology, Wuhan, Hubei 430065, China;

2. Key Laboratory of Quark and Lepton Physics (MOE) and Institute of Particle Physics, Central China Normal University, Wuhan 430079, China)

Abstract: Artificial Neural Network (ANN) has become a powerful tool in the field of scientific research with its powerful information encapsulation ability and convenient variational optimization method. In particular, there have been many recent advances in computational physics to solve variational problems. Deep neural network (DNN) is used to represent the wave function to solve quantum many-body problems using variational optimization. In this work we used a new Physics-Informed Neural Network (PINN) to represent the Cumulative Distribution Function (CDF) of some classical problems in quantum mechanics and to obtain their ground state wave function and ground state energy through the CDF. By benchmarking against the exact solution, the error of the results can be controlled at a very low level. This new network architecture and optimization method can provide a new choice for solving quantum many-body problems.

Key words: deep neural network; variational problems; the Cumulative distribution function; ground state wave function

1　Introduction

With the rapid growth of data and computing resources, Machine Learning(ML) technology is playing an increasingly important role in more and more areas. For some physical phenomena expressed by partial differential equations，it is difficult for traditional methods to solve complex situations such as high dimensions. Universal approximation theorem[1] enables deep neural networks to solve variational function problems. The deep learning based method proposed in this paper is a new solution, by minimizing the difference between variational functions and partial differential equations, DNN can transform physical problems[2] into optimization problems[3] that can be solved with modern deep learning libraries.

In the case of variational equations, physical constraints play an essential role[4-8]. In this paper, instead of using DNN to represent the wave function directly, we use DNN to represent the cumulative distribution function $\int_{-\infty}^{x} f(x')dx'$ of the wave function $\psi(x)$. In order to improve the efficiency of data learning, we apply the unitary constraint to the variational function, and the whole computational process does not need to compute numerical integration, thus greatly reducing the computational amount. These advantages make the method in this paper more suitable for dealing with multiple problems with a large amount of computation.

In this paper, we reviewed that physically constrained neural networks can be used as efficient trial wave functions in solving Schrodinger equations with the harmonic oscillator and the Woods-Saxon. We have also extended this method to Schrodinger equation with the double-well potentials in the same framework.

2　Methods

Unlike the previous direct expression of the wave function by the Variational Artificial Neural Networks (VANN)[5,9-11], we obtain the wave function indirectly

through the CDF function, and the relationship between the two is as follows

$$F(x) = \int_{-\infty}^{x} \psi^*(x')\psi(x')dx' \tag{1}$$

$$\psi(x) = \sqrt{\frac{dF(x)}{dx}} \tag{2}$$

where $F(x)$ is the CDF，$\psi(x)$ is the wave function. Auto differentiation (auto-diff)[12] tools in deep learning libraries can be used to compute the derivative dF/dx. The wave function obtained by the CDF satisfies the normalisation, thus avoiding numerical integration and simplifying the calculation. Furthermore, the value range of the CDF is between $(0, 1)$, which makes it much easier to train. Due to the above two advantages, our algorithm has less computation and training cycles, while maintaining good accuracy.

The input of the neural network is the n-dimensional spatial coordinates x. The first layer of the DNN has $m = 32$ hidden neurons whose values are computed by $h_1 = \sigma(xW_1 + b_1)$, where W_1 is the weight matrix with $n \times m$ elements and b_1 is the bias vector with m values. To avoid linear superposition, a non-linear function σ, called the activation function, is introduced, which can improve the representational ability of the DNN. The second layer works similarly, taking the output of the first layer as the input of the second layer $h_2 = \sigma(h_1 W_2 + b_2)$, allowing multiple hidden layers to be stacked. The last layer returns the value of the CDF function.

Since the CDF function is monotonic, the variational function of the DNN fitting must also be monotonic. To solve this problem, we add a non-negative constraint to the weight W_i. Theoretically, nonlinear functions such as sigmoid, tanh and relu can all be used as the activation function of the DNN, but in practice, since we need continuous and monotonic activation functions, and the derivative cannot be 0, we finally choose the sigmoid function as the activation function, which satisfies the above conditions. And the sigmoid function has a range of $(0, 1)$,which guarantees the output range of the last layer.

The goal of the training is to minimise the violation between the wave function obtained by the DNN and the Schrodinger equation, so that the ground state energy and the corresponding wave function can be found.

$$H|\psi\rangle = E_0|\psi\rangle \tag{3}$$

where $H = -\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{x})$ is the Hamiltonian operator and E_0 represents its smallest eigenvalue. The loss function is defined as follow,

$$L(\theta) = |(H - E_0)|\psi\rangle + |F(x_{min})| + |F(x_{max}) - 1| \tag{4}$$

where θ is a trainable parameter, including W_i, b_i and E_0, E_0 is another trainable parameter initialised to 0. DNN guarantees that ψ satisfies the Schrodinger equation by minimising the first term, and the other two terms guarantee that the CDF function fitted by DNN satisfies the boundary conditions. The range of the CDF provides normalisation of the wave function, which simplifies the calculation.

Note that the initial values of the trainable parameters W_i and b_i are usually initialised randomly or in the Xavier scheme[13]. In our work, however, the initialisation scheme of the parameter has little effect on the result, so we use random initialisation.

We add another loss term $e^{\gamma E_0}$, where γ is a small value. By subtracting this term from the loss function, integration is avoided and DNN optimizes the loss function to make E_0 lower. By training this energy term, the ground state energy and the ground state wave function can be obtained with small error[14] .

We tested the performance of the DNN Schrodinger equation solver on three classical quantum mechanical problems. The first problem is the harmonic oscillator model[15], the potential of the harmonic oscillator can be written as,

$$V = \frac{1}{2}m\omega^2 x^2 \tag{5}$$

where m is the mass of the oscillator, ω is its angular frequency and x is its deviation from equilibrium position.

The second example is to solve Schrodinger equation with Woods-Saxon potential[16],

$$V = \frac{-1}{1 + e^{\frac{|x|-R_0}{a_0}}} \tag{6}$$

where a_0 is the thickness of the surface layer and the average radius is given as R_0.

The third is a double-well potential problem, which is a new potential added to our previous work[14]. The potential in the Hamiltonian can be written as:

$$V = -e^{-h(x+l-a)(x+l+a)} - e^{-h(x-l+a)(x-l-a)} \tag{7}$$

Where h,l,a respectively control the depth, position, and width of the double-well. Unlike the first two questions,the ground state wave function of a double-well potential problem will exhibit two peaks. Applying our method to such a potential can further verify the universality of the architecture.

For simplicity, some parameters are set as follows,

$$\hbar = m = \omega = 1, R_0 = 6.2, a_0 = 0.1, h = 4, a = 0.5, l = 1.5 \tag{8}$$

The input to DNN is an unordered list of domain coordinates. The ground state energy and the CDF function are obtained by minimising the terms of the loss function, and thus the ground state wave function is obtained indirectly. Unlike the previous method of solving the Schrodinger equation using supervised learning, our method belongs to unsupervised learning. The wave function is usually very close to the exact wave function after training the DNN through iterations in harmonic oscillator, Woods-Saxon potential, infinitely high potential well[14] and a double-well potential. DNN can get very accurate results after 2000 iterations, but we trained 10,000 times, using a very small learning rate in the last 1,000 times.

We used Adam algorithm[14,17-18] that add momentum mechanism and adaptive learning rate to the simple stochastic gradient descent,

$$\theta_{n+1} = \theta_n - lr \frac{1}{m} \sum_{i=1}^{m} \frac{\partial l_i}{\partial \theta} \tag{9}$$

where θ represents the model parameters, including weights and biases within a neural network, The term n denotes the current iteration number, reflecting the step in the iterative process of optimization. 'lr', the learning rate, is a critical hyperparameter that controls the size of the steps taken towards the minimum of the loss function. m is the batch size, which determines the number of samples used for each parameter update. $\frac{\partial l_i}{\partial \theta}$ represents the gradient of the loss function l_i with respect to θ, guiding the direction and magnitude of parameter adjustments.

We use a learning rate scheduler to speed up the training, setting a higher learning rate in the early stages to speed up convergence, and a lower learning rate in the later stages to reach the global minimum.

To quantify the difference between the wave function obtained by DNN ψ_{DNN} and the real wave function ψ_{true}, we introduce the partial wave fidelity K,

$$K = \frac{< \psi_{true} | \psi_{DNN} >^2}{< \psi_{true} | \psi_{true} >< \psi_{DNN} | \psi_{DNN} >} \tag{10}$$

The closer the K is to one, the closer the result of DNN is to the exact wave function.

3   Results

For the harmonic oscillator problem, DNN can control the ground state energy error within 0.06% after 1500 training, and reduce the error to 0.002% after 10000 training, the partial wave fidelity K is 0.9999967 when using 4 hidden layers with 16 units per layer. To study the influence of the number of variational parameters on the training results, K is discussed in Ref.[14] by using different numbers of hidden layers and units per layer.

To understand the difference between the wave function obtained by DNN and the real wave function, we calculate the CDF function, the wave function and the first and second derivatives of the wave function respectively.
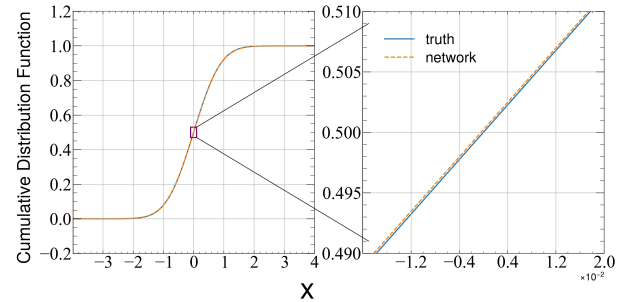

Fig. 1   One-dimensional harmonic oscillator's CDF function[14].(The image on the right is a partial enlargement of the image on the left)

The error between the CDF function fitted by DNN and the real CDF function is within 0.0001 as shown in Fig. 1. The error range of the ground state wave function can be controlled within 0.0002 as shown in Fig. 2. In addition, the accuracy of the first and second derivatives of the wave function obtained by the DNN is also very high, which means that the DNN has learned the real physical information of the wave function.

We verified the universality of the method by using the same network to solve the Schrodinger equation with the Woods-Saxon potential and the double-well potential. For the Schrodinger equation with Woods-Saxon potential energy, using the same parameter γ in the harmonic oscillator potential, the Woods-Saxon potential could control the error
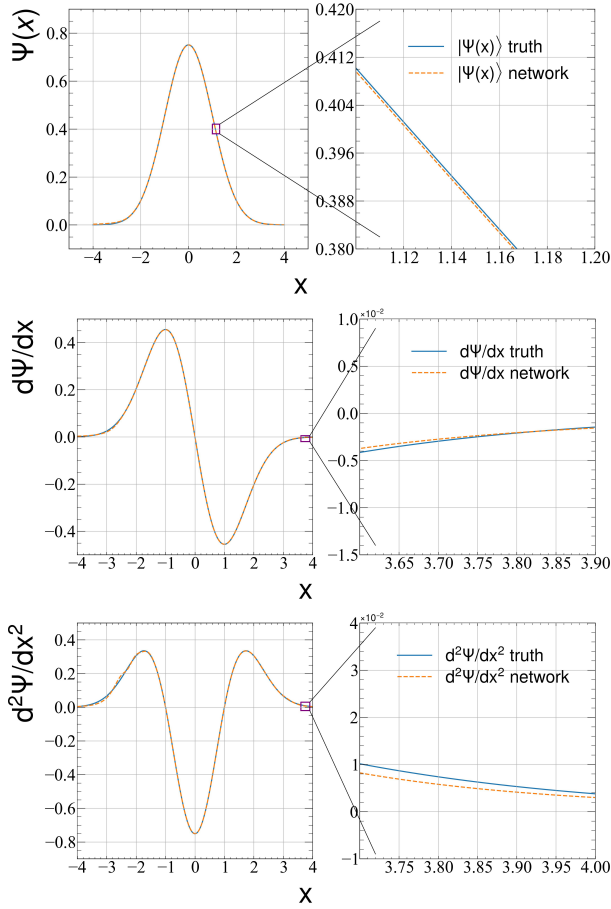
Fig. 2　One-dimensional harmonic oscillator's ground state wave function(top panel), and it's first derivative(central) and the second derivative(bottom)[14].
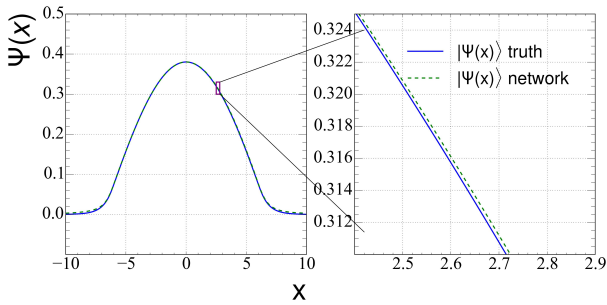

Fig. 3　The ground state wave function in the Woods-Saxon potential energy[14].
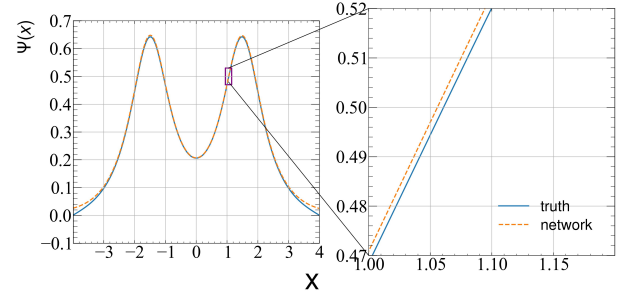

Fig. 4　The ground state wave function double-well potential.

of the ground state energy less than 0.02%. Fig. 3 shows the comparisons between the learned ground state wave function and the true values. It can be seen that the ground state wave function obtained from the DNN also agrees very well with the exact solution. The error range can be controlled within 0.0002 for the ground state wave function of the Woods-Saxon potential as shown in Fig.3. And the ground state energy calculated by the DNN for the Woods-Saxon potential problem is -0.97382, which is also within an error of 0.002% to the exact result of -0.97385 and $K = 0.999964$.

However, compared with the results of harmonic oscillator and Woods-Saxon potential, because the ground state energy is too close to the first excitation state energy under the double-well potential, we need to appropriately increase the energy layer parameters γ so that DNN gives priority to energy optimization, in which case the ground state energy error of the double-well potential is less than 0.5%. In addition, we make the DNN more expressive by increasing the number of hidden layers, which brings the value of K closer to 1. The relative error of the ground state energy can be kept within 0.5% and K remains stable around 0.99734 for the double-well potential problem. It will still maintain a high level of accuracy, although it will not be able to reach the accuracy of the first two problems.

4　Conclusions

We use DNN to fit the CDF function of the ground state wave function, so that the resulting wave function automatically satisfies the normalisation and simplifies the computation. By minimising the violation of the trial wave function and the trial ground state energy E_0 to the Schrodinger equations, the variational problem is reduced to an optimization problem. This method improves the accuracy of the Schrodinger equation solution with less computation. The effectiveness of this method is confirmed by solving the harmonic oscillator, the Woods-Saxon potential and the double-

well potential with small relative error.

$E_0 = \frac{\langle\psi|H|\psi\rangle}{\langle\psi|\psi\rangle}$ where numerical integration is required for both the numerator and the denominator in traditional variational methods for solving quantum mechanical problems. In another DNN Schrodinger solver[9-10], the form of a trial wave function should be approximated to the exact solution for eliminating fluctuation and the initial values of the network parameters also greatly affect the optimization results. In our framework, the goal is to minimise the violation of the Schrodinger equation on the sampled spatial coordinates. The trial wave function is automatically normalized based on the constraints of the constructed neural network. The normalized ground wave function improves the efficiency of the E_0 calculation with a lower relative error. Most importantly, our algorithm is more universal and provides the possibility to solve problems that have never been solved before, because it can directly ignore the pre-training process and does not need to know any information about the exact solution before training.

There are a number of improvements that can be made to the current method. For example, active learning or coordinate sampling by the learned wave function can be used to improve training efficiency. We could improve the accuracy in solving the double-well potential problem by optimising the structure of our framework. The DNN framework will be optimized by solving the Schrodinger equation with a 3-dimensional potential. It may also open up a new way to solve many nucleon problems.

5　Acknowledgments

References

[1] CYBENKO G V. Mathematics of Control, Signals and Systems, 1989, 2: 303. DOI: 10.1007/BF02551274.

[2] BOEHNLEIN A, et al. Rev Mod Phys, 2022, 94(3): 031003. DOI: 10.1103/RevModPhys.94.031003.

[3] SAAD D. American Scientist, 2004, 92(6): 578.

[4] HENDRIKS J, JIDLING C, WILLS A, et al. arXiv preprint arXiv:200201600, 2020. DOI: https://doi.org/10.48550.

[5] HERMANN J, SCHÄTZLE Z, NOÉ F. Nature Chemistry, 2020, 12 (10): 891. DOI: https://doi.org/10.1038/s41557-020-0544-y.

[6] PFAU D, SPENCER J S, MATTHEWS A G D G, et al. Phys Rev Res, 2020, 2: 033429. https://link.aps.org/doi/10.1103/PhysRevResearch .2.033429.

[7] LUO D, CARLEO G, CLARK B K, et al. Phys Rev Lett, 2021, 127: 276402. https://link.aps.org/doi/10.1103/PhysRevLett.127.276402.

[8] ABBOTT R, ALBERGO M S, BOTEV A, et al. arXiv preprint arXiv:220803832, 2022. DOI: 10.48550/arXiv.2208.03832.

[9] KEEBLE J, RIOS A. Phys Lett B, 2020, 809: 135743. DOI: https: //doi.org/10.1016/j.physletb.2020.135743.

[10] SAITO H. Journal of the Physical Society of Japan, 2018, 87(7): 074002. DOI: https://doi.org/10.7566/JPSJ.87.074002.

[11] CARLEO, GIUSEPPE, TROYER, et al. Science, 2017. DOI: 10.112 6/science.aag2302.

[12] PASZKE A, GROSS S, CHINTALA S, et al. Automatic differentiation in pytorch[Z]. 2017.

[13] GLOROT X, BENGIO Y. Understanding the difficulty of training deep feedforward neural networks[C]//Proceedings of the thirteenth international conference on artificial intelligence and statistics. JMLR Workshop and Conference Proceedings, 2010: 249.

[14] PU K, LI H, LU H, et al. Chin Phys C, 2023. DOI: https://doi.org/10 .1088/1674-1137/acc518.

[15] SENITZKY I. Physical Review, 1961, 124(3): 642. DOI: 10.1103/ PhysRev.124.642.

[16] CAPAK M, GÖNÜL B. Modern Physics Letters A, 2016, 31(23): 1650134. DOI: https://doi.org/10.48550/arXiv.1607.02742.

[17] ABADI M, AGARWAL A, BARHAM P, et al. arXiv preprint arXiv:160304467, 2016. DOI: https://doi.org/10.48550/arXiv.1603. 04467, Focus, to, learn, more.

[18] KINGMA D P, BA J. arXiv preprint arXiv:14126980, 2014. DOI: https://doi.org/10.48550/arXiv.1412.6980, Focus, to, learn, more.

用软约束单调神经网络求解薛定谔方程

李汉林[1]，刘 轩[1]，蒲开放[1]，庞龙刚[2]

(1. 理学院，武汉科技大学, 湖北省武汉市 430065; ;
2. 华中师范大学粒子物理研究所，夸克与轻子物理教育部重点实验室，湖北省武汉市 430079)

摘要： 人工神经网络 (Artificial Neural Network, ANN) 以其强大的信息封装能力和方便的变分优化方法成为科学研究领域的有力工具。特别是最近在计算物理解决变分问题方面取得了许多进展。用深度神经网络 (DNN) 表示波函数来求解变分优化的量子多体问题时，我们使用一种新的物理信息神经网络 (PINN) 来表示量子力学中一些经典问题的累积分布函数 (CDF)，并通过 CDF 获得他们的基态波函数和基态能量。通过对精确解的基准测试，可以将结果的误差控制在很小范围。这种新的网络结构和优化方法为解决量子多体问题提供了新的选择。

关键词： 深度神经网络; 变分问题; 累积分布函数; 基态波函数